



TITLE:

INFORMATION THEORY AND STATISTICAL MECHANICS OF RANDOM MATRICES

AUTHOR(S):

Hasegawa, Hiroshi

CITATION:

Hasegawa, Hiroshi. INFORMATION THEORY AND STATISTICAL MECHANICS OF RANDOM MATRICES. 数理解析研究所講究録 1998, 1066: 205-220

ISSUE DATE:

1998-10

URL:

<http://hdl.handle.net/2433/62473>

RIGHT:

INFORMATION THEORY AND STATISTICAL MECHANICS OF RANDOM MATRICES

Hiroshi Hasegawa

*Research Center of Quantum Communications, Tamagawa University,
Machida, Tokyo 194, Japan*

Abstract. Balian's program of assigning a probability distribution to a random matrix is exploited to construct a possible general form of the joint eigenvalue distribution for an ensemble of hermitians or unitaries. From a point of view of open system, it can be viewed as the equilibrium statistical mechanics of an eigenvalue system in contact with Gaussian environmental reservoir. The statistical aspect of the previously introduced Hamiltonian level dynamics is discussed by putting it into this framework and is illustrated by way of a concrete, exact result of Gaudin's model in 1966.

1. On Balian's Article in 1968

As an introduction of the present review on *an aspect of Random Matrix Theories*, the author wishes to begin with a brief comment about Balian's work some thirty years ago[1], which implemented an information theoretical approach to random matrix studies for the first time. The decade of 1960's was the initial development of this field, where the name of Wigner, Dyson, Porter, Mehta and others may be cited. Its revise after 1980's was stimulated in connection with broad interests in *quantum chaos studies* and applications to various branches of Physics and Chemistry: it may be referred to a recent comprehensive review in Physics Reports [2]. However, even in this report, little has been mentioned about information theoretical viewpoint(the same feature can be said also about Mehta's book[3]).

In the present article, we would like to emphasize the important role of information theory to play in the basic part of constructing random matrix theories along the proposal by Balian cited above: this is expressed in his prescription summarized by two postulates concerning *information quantity on a continuous probability space* as follows. Let $P[M]$ denote a probability distribution to be assigned to a sample matrix M taken up from an ensemble of matrices, and suppose that one wishes to know $P[M]$. Write the information quantity associated with the ensemble as

$$I\{P[M]\} \equiv \int d[M] P[M] \log P[M]. \quad (1.1)$$

Then, *postulate A*: a (Riemannian) metric is defined, in the matrix space to which M belongs, by the expression

$$ds^2 \equiv \text{tr} \delta M \delta M^*. \quad (1.2)$$

postulate B: among all probability distributions $P[M]$ for a random matrix constrained to satisfy some given properties, one should adopt the distribution which minimizes $I\{P[M]\}$,

and is thus the least biased (equivalent to the *maximum entropy principle under pertinent constraints*).

Therefore, it would be worthwhile to focus on how these two postulates act in the actual formalism of random matrix theories in the recent as well as old development, and the present article is written to satisfy this purpose. Specifically, the author wishes to put his emphasis on the role of *Riemannian geometry in statistics* implied by postulate A which Balian supposedly intended to clarify.

2. Information Theory and Statistical Mechanics of Gaussian-Reservoir Open Systems

2.1. Preliminary[4]

Suppose that we are given two sets of random variables, X and Y which let us call *object variable* for X and *reservoir variable* for Y , respectively. The former variable describes our object system, while the latter an environmental system against the former. Such a situation often occurs in statistical physics, and here we aim at a formulation of the special situations that the reservoir system may be subject to a Gaussian distribution.

A Gaussian distribution, very familiar to every scientists, can be written only in terms of two kinds of parameter i.e. the *mean* and the *variance* of the random variable, and for a d -dimensional vector case it can be written as

$$P_G(y) = \frac{1}{((2\pi)^d |V|)^{1/2}} \exp\left[-\frac{1}{2}(y - \langle y \rangle)' V^{-1} (y - \langle y \rangle)\right], \quad y \in \mathbb{R}^d, \quad (2.1)$$

where $\langle y \rangle$ represents the expectation(mean) value of the random vector variable y denoted by $\text{mean}(Y)$ and $'$ implies the transpose of a vector. In the exponential part of the above expression, V represents a positive, symmetric $d \times d$ tensor, and is related to the expectation value

$$V_{\mu,\nu} = \langle (y - \langle y \rangle)_\mu (y - \langle y \rangle)_\nu' \rangle \quad (2.2)$$

which is called the covariance of Y denoted by $\text{Cov}(Y, Y)$. Also, $|V|$ denotes the determinant of the tensor V which is assumed to be positive-definite.

In the present studies of open systems, we must incorporate the object variable X into the above distribution, and most generally $\text{mean}(Y)$ and $\text{Cov}(Y, Y)$ must be regarded as functions of X which is assumed to be an n -dimensional vector random variable. Then, in order to be able to write the joint distribution for both variables X and Y , the Gaussian distribution (2.1) must be modified such that its exponential part also represents a probability distribution on the object variable X . Let us redefine y by means of its shift, $y - \langle y \rangle$, in the exponential, and write

$$P_G(x, y) = \frac{1}{Z} \exp\left[-\frac{1}{2} y' V(x)^{-1} y\right], \quad y \in \mathbb{R}^d, x \in D, \quad (2.3)$$

where

$$Z = \int_D \int_{\mathbb{R}^d} dx dy \exp\left[-\frac{1}{2} y' V(x)^{-1} y\right] = (2\pi)^{d/2} \int_D |V(x)|^{1/2} dx \quad (2.3a)$$

is assumed to exist, that is, the integral $\int_D |V(x)|^{1/2} dx$ is assumed to be finite.

We define the conditional probability density function of the Y -variable, when the X is conditioned by fixing its value $X = x$, by the formula

$$P_G(x, y) \equiv P_G(y|X = x) \cdot P(x) \quad (2.4)$$

under normalization condition

$$\int_{\mathbb{R}^d} P_G(y|X = x) dy = 1, \quad \text{and} \quad \int_D P(x) dx = 1. \quad (2.5)$$

This yields the expression for $P(x)$, namely

$$P(x) = \int_{\mathbb{R}^d} P_G(x, y) dy = \frac{(2\pi)^{d/2}}{Z} |V(x)|^{1/2} = \frac{|V(x)|^{1/2}}{\int_D |V(x)|^{1/2}}, \quad (2.6)$$

called the *reduced* (or, *coarse-grained*) probability distribution, and also for $P_G(y|X = x)$

$$P_G(y|X = x) = P_G(x, y)/P(x) = \text{eq. (2.1) with } V = V(x),$$

identifying that the original Gaussian distribution is the conditional probability density function of the Y -variable for each fixed x . Thus,

$$P_G(y|X = x) = \frac{1}{((2\pi)^d |V(x)|)^{1/2}} \exp[-\frac{1}{2} y' V(x)^{-1} y], \quad y \in \mathbb{R}^d, \quad x \in D. \quad (2.7)$$

2.2. Entropy and conditional entropy

We adopt the standard definition of *entropy* for a continuous random variable $X \in D \subseteq \mathbb{R}^d$ with probability density function $p(x)$:

$$H(X) \equiv h(p) = \int_D p(x) [-\log p(x)] dx \quad (2.8)$$

(the so-called differential entropy of Shannon)

Unlike the entropy on a discrete probability space, $H(X)$ may take negative values, but except with this point, the above definition of entropy satisfies most of the basic properties which are known for the concept in the field of *information theory*, and we list some of them: (1) for any two probability density functions $p(x)$ and $q(x)$,

$$H(X) = \int_D p(x) [-\log p(x)] dx \leq \int_D p(x) [-\log q(x)] dx, \quad (2.9)$$

where the equality holds only when $q(x) = p(x)$ identically (Kullback inequality [4]).

(2) $H(X, Y)$ is defined as the entropy for the joint probability density $p(x, y)$, and $H(Y|X)$ as the mean of the conditional probability density $p(y|X = x)$ over the reduced probability density $p(x) = \int p(x, y) dy$ so that $H(Y|X) = \int_D H(Y|X = x) p(x) dx$. Then, the following two identities hold.

$$H(X, Y) = H(Y|X) + H(X) = H(X|Y) + H(Y). \quad (2.10)$$

(3) Let $H(X)$ and $H(Y)$ be the entropy of the reduced probability density $p(x)$ and $p(y)$ of a joint probability density $p(x, y)$, respectively. Then,

$$H(X, Y) \leq H(X) + H(Y), \quad H(X|Y) \leq H(X) \text{ and } H(Y|X) \leq H(Y). \quad (2.11)$$

These inequalities indicate that a coarse-graining generally gives rise to an increase of the entropy from its value before the coarse-graining.

Let us now apply these results to our Gaussian-reservoir open system:

$$P_G(x, y) = \frac{1}{Z} \exp[-\frac{1}{2} y' V(x)^{-1} y], \quad Z = (2\pi)^{d/2} \int_D |V(x)| dx;$$

$$H(X, Y) = \log Z + \frac{d}{2} = \log(2\pi e)^{d/2} + \log \int_D |V(x)|^{1/2} dx, \quad (2.12)$$

$$P_G(y|X=x) = \frac{1}{((2\pi)^d |V(x)|)^{1/2}} \exp[-\frac{1}{2} y' V(x)^{-1} y];$$

$$H(Y|X=x) = \log P(x) + \log Z + \frac{d}{2} \quad (2.13)$$

$$H(Y|X) = -H(X) + \log Z + \frac{d}{2} \quad (2.14)$$

$$P(x) = \int_D P_G(x, y) dy = \frac{|V(x)|^{1/2}}{\int_D |V(x)|^{1/2} dx}, \quad \text{also by virtue of eqs. (2.13) and (2.12)}$$

$$= \frac{1}{(2\pi)^{d/2} \int_D |V(x)|^{1/2} dx} e^{H(Y|X=x)}, \quad (2.15)$$

$$H(X) = \int_D P(x) [-\log P(x)] dx = -H(Y|X) + H(X, Y). \quad (2.16)$$

The last expression (2.16), which verifies the first of two identities in (2.10), may yield a thermodynamic interpretation that the reduced-system entropy can be expressed as

$$H(X) = (\text{inverse temperature}) \times [\text{energy} - \text{free energy}], \quad (2.17)$$

where $-H(Y|X)$ corresponds to *energy* and $-H(X, Y)$ to *free energy*, respectively, of the object system X . This is characteristic to the present Gaussian-reservoir open system.

2.3. Maximum conditional-entropy principle

The well-known property of the Gaussian distribution $p_G(y)$ is that among all distributions of $y \in \mathbb{R}^d$ with mean 0 and covariance V (fixed), it has the unique maximum value of $H(Y)$ which is equal to $\log((2\pi e)^d |V|)^{1/2}$. The proof of this fact can best be seen by the application of Kullback inequality (2.9) to $q(y) = p_G(y)$ whose logarithm is a quadratic function of the y variable, yielding the expectation value V when averaged over the other arbitrary distribution $p(y)$, and by the fact that the equality holds only when $p(y) = q(y)$. We may express this property as

$$\max_p H(Y) (= \max_p h(p)) = H_G(Y) (= h(p_G))$$

$$\text{under constraint mean}(Y) = 0, \text{ and Cov}(Y, Y) = V,$$

and apply the same principle to the reservoir system Y for each given configuration of the object system $X = x$. Thus,

$$\max_{p(y|X=x)} H(Y|X=x) = H_G(Y|X=x) \equiv h_{\max}(p(y|X=x)) \quad (2.18)$$

under constraint

$$\text{mean}(Y) = 0, \quad \text{and} \quad \text{Cov}(Y, Y) = V(x). \quad (2.18a)$$

The result of the maximization can be represented as the expression(see eq.(2.15)) for $P(x)$:

$$P(x) = \frac{1}{(2\pi)^{d/2} \int_D |V(x)|^{1/2} dx} e^{h_{\max}(p(y|X=x))}. \quad (2.19)$$

This form of *maximum entropy principle* is different from the usual variational principle in statistical mechanics, and the relation between them should be clarified by invoking the thermodynamic interpretation (2.17), where apart from numerical factors,

$$H_G(X, Y)(\text{total entropy after maximization}) = (\text{free energy of object system}). \quad (2.20).$$

3. Possible Riemannian Metrics and Distributions on Random Matrix Spaces

In order to equip an ensemble of matrices with a Gaussian-reservoir structure discussed in the preceding section, we must introduce a Riemannian metric into the space of matrices, and for this purpose let us recall Balian's *postulate A* (1.2) concerning the distance between two infinitesimally separated matrices. A Riemannian metric tensor $(g_{\mu,\nu})$ can then be defined as the coefficient tensor of the distance ds^2 with respect to a quadratic form of an infinitesimal parameter set. For this purpose, a recent paper by Petz[5] provides a suggestive and useful idea, proposing an axiomatic approach to possible Riemannian metrics on matrix spaces.

Let us denote, following Petz, the space of $n \times n$ complex matrices by \mathcal{M}_n on which a sesqui-linear form $K(B, A)$ (linear with respect to A and anti-linear to B ; $A, B \in \mathcal{M}_n$) is defined. The simplest example of such is the Hilbert-Schmidt inner product $K_{H-S}(B, A) \equiv \text{tr} B^* A$ which satisfies the unitary invariance, namely

$$K(U^* B U, U^* A U) = K(B, A), \quad (3.1)$$

and, importantly, it is known that such a unitary invariant sesqui-linear form must be identical to the H-S inner product.

However, if the form K is supposed to depend on another hermitian matrix, then there may arise a variety of inner products of non-HS type, and it is desirable to classify such inner products as legitimate forms. Denoting the set of all hermitian matrices in \mathcal{M}_n by \mathcal{M}_n^s , we list up the properties of the expected forms as follows.

- (a)symmetry $K_H(A^*, B^*) = K_H(B, A)$; $H \in \mathcal{M}_n^s$, $A, B \in \mathcal{M}_n$. When A and B are restricted to hermitians, the form K becomes real and symmetric, and hence it is a bilinear form.
- (b)positive definiteness $K_H(A, A) \geq 0$, and the equality holds only when $A = 0$.
- (c)continuity of the map $H \mapsto K_H$: this continuity holds for every $A \in \mathcal{M}_n$ in $K_H(A, A)$.
- (d')unitary covariance $K_{U^* H U}(U^* B U, U^* A U) = K_H(B, A)$: this relaxes the condition of unitary invariance in the strict sense to the same condition but with an inclusion of the subsidiary matrix H , and hence the bilinear form K_H belongs to much wider class than the Hilbert-Schmidt inner product.

This last condition (d') is important for the present purpose, and actually is weaker than the condition (d) of *monotonicity* which Petz has proposed, setting it up for a density matrix D . (A density matrix D in \mathcal{M}_n is a special hermitian matrix, positive and $\text{tr} D = 1$.)

(d)monotonicity $K_{T(D)}(T(A), T(A)) \leq K_D(A, A)$, where T , a super-operator (a linear map) $\mathcal{M}_n \mapsto \mathcal{M}_m$, in which a positive matrix is mapped to a positive matrix (called *stochastic map*).

An intuitive understanding of the monotonicity of T is that by any coarse-graining of the pertaining matrices in K_D , i.e. both A and D , the metric represented by K_D must be a non-increasing quantity. When T is a unitary map, the above monotonicity inequality becomes the equality, because now T can be an invertible super-operator from \mathcal{M}_n onto itself. Therefore, condition (d) includes (d') (d) is more stringent than (d'): if (d) is valid for a form K , (d') is also valid for the same form, but the converse is not necessarily true).

The condition (d) (or (d')) enables one to take the representation of the pertinent matrices where D (or H) is diagonal, and to exhibit the form of K in terms of the matrix elements A_{jk} : Petz has shown that, under condition (d) with $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$

$$K_D(A, A) = \sum_{j,k} c(\lambda_j, \lambda_k) |A_{jk}|^2, \quad (3.2)$$

where the real function $c(\lambda, \mu)$ satisfies that

$$c(\lambda, \mu) = c(\mu, \lambda), \quad c(\lambda, \lambda) = 1/\lambda, \quad c(t\lambda, t\mu) = t^{-1}c(\lambda, \mu). \quad (3.3)$$

Thus, only a single, continuous function $c(x)$ is enough to represent a monotone metric on a matrix space, as far as the dimensionality is finite, which is related to an operator-monotone function according to Petz[5], and is used to characterize a quantum mechanical Fisher metric[6].

We will seek the same kind of representation of $K_H(A, A)$ under condition (d'). For this purpose, let us adopt another condition (d''):

(d'')translational invariance with respect to H $K_{H+aI}(B, A) = K_H(B, A)$.

It is straightforward to show that, under conditions (d') and (d'') with $H = \text{diag}(\lambda_1, \dots, \lambda_n)$

$$K_H(A, A) = \sum_{j \leq k} c(\lambda_j, \lambda_k) |A_{jk}|^2, \quad (3.4)$$

where the real function $c(\lambda, \mu)$ satisfies that

$$c(\lambda, \mu) = c(\lambda - \mu) > 0 \quad \lambda \neq \mu \quad \text{and} \quad c(\lambda, \lambda) = (\text{independent of } \lambda) \geq 0. \quad (3.5)$$

We have just obtained a general form of Riemannian metric on a matrix space \mathcal{M}_n under conditions (a) \sim (d') and (d''), where the quadratic quantity $|A_{jk}|^2$ indexed by $\frac{1}{2}n(n+1)$

pairs (j, k) represents the square of a *tangent vector* component.

We aim at a Gaussian-reservoir distribution on the matrix space \mathcal{M}_n by means of this metric with a $d(= \frac{1}{2}n(n+1))$ -dimensional (complex) tangent vector. For this purpose, let us go to the *cotangent vector* defined by

$$y_{j,k} \equiv c(\lambda_j - \lambda_k)A_{jk} \quad j \neq k; \quad y_{j,j} \equiv 0 \quad (c(0) = 0 \text{ assumed}). \quad (3.6)$$

Then,

$$K_H(A(y), A(y)) = \sum_{j < k} \frac{1}{c(\lambda_j - \lambda_k)} |y_{j,k}|^2, \quad (3.7)$$

in which λ represents an n -dimensional object variable, and y a d -dimensional reservoir variable.

Let us change the notation of our object variable to conform to Sec.2 and write the result for the present distribution for random matrices:

$$P_G(x, y) = \frac{1}{Z} \exp\left[-\frac{1}{2} \sum_{j < k} \frac{1}{c(x_j - x_k)} |y_{j,k}|^2\right] \quad (3.8)$$

$$\text{mean}(Y) = 0, \quad \text{Cov}(Y, Y) = \text{diag}(\dots, c(x_j - x_k), \dots) \quad (3.8a)$$

$$\text{i.e.} \quad \langle y_{j,k} y_{m,n} \rangle = c(x_j - x_k) \quad \text{for } (j, k) = (m, n); \quad = 0 \text{ for } (j, k) \neq (m, n).$$

In the matrix space \mathcal{M}_n , we introduce the matrix random variable Y whose elementary units are the set of all matrix units $(e_{j,k})$, writing

$$Y = \sum_{j,k} e_{j,k} y_{j,k} \quad (y_{k,j} = y_{j,k}^*). \quad (3.9)$$

Then, on the basis of the maximum conditional-entropy principle under constraint (3.8a), the resulting Gaussian distribution (3.8) expresses the following properties which are remarkable.

(i) statistical independence of different units

$$\text{for } (j, k) \neq (m, n), \quad P(y_{j,k}, y_{m,n}) = P(y_{j,k}) \cdot P(y_{m,n}). \quad (3.10)$$

(ii) identical distribution for all the units with off-diagonal type

$$\text{Cov}(Y_{j,k}^*, Y_{j,k}) \text{ depends on the pair } (j, k) \text{ only through } x_j - x_k \text{ in a common function } c(.). \quad (3.11)$$

Consequently, the determinant $|V|$ of the covariance function $V(x)$ is simply the product of all the pairs of single variance $c(x_j - x_k)$, and we can get the reduced probability distribution for the object eigenvalue system as follows:

$$P(x_1, x_2, \dots, x_n) = C_n \prod_{j < k} c(x_j - x_k)^{\nu/2}, \quad (3.12)$$

where,

$$C_n = [\int_D \prod_{j < k} c(x_j - x_k)^{\nu/2} dx_1 \dots dx_n]^{-1} \quad \nu = 1, 2 \text{ and } 4, \quad (3.12a)$$

the integer ν being the multiplicity of the real component of each cotangent vector $y_{j,k} (j \neq k)$ i. e. $\nu = 1$ for real part only, $\nu = 2$ for real and imaginary parts, and $\nu = 4$ for four quaternion parts if Pauli spins are involved. Also, by regarding this index ν as a continuous parameter of *inverse temperature*, and apart from the pure numerical factor $\log(2\pi e)^{d/2}$,

$$\text{free energy of the reduced system } -\frac{1}{\nu} H(X, Y) = \frac{1}{\nu} \log \int_D \prod_{j < k} c(x_j - x_k)^{\nu/2} dx_1 \dots dx_n \quad (3.13)$$

$$\text{energy of the reduced system } -\frac{1}{\nu} H(X|Y) = n \times \sum_{1 < k} \langle \log c(x_1 - x_k)^{1/2} \rangle_{P(x)}. \quad (3.14)$$

4. Standard Random Matrix Distributions

joint N-revel distribution; Gaussian ensembles

$$P_N(x_1, x_2, \dots, x_N) = C_{N\beta} e^{-\frac{\beta}{2}(x_1^2 + \dots + x_N^2)} \prod_{j < k} (x_j - x_k)^\beta, \quad -\infty < x_j < \infty \quad (4.1)$$

$$C_{N\beta} = (\text{see ref. [3]}); \beta = 1 (GOE), = 2 (GUE), = 4 (GSE) \quad (4.1a)$$

joint N-level distribution; circular ensembles

$$P_N(x_1, x_2, \dots, x_N) = C'_{N\beta} \prod_{j < k} |e^{i(x_j - x_k)} - 1|^\beta, \quad -\pi < x_j < \pi \quad (4.2)$$

$$C'_{N\beta} = \frac{(\Gamma(1 + \beta/2))^N}{(2\pi)^N \Gamma(1 + \beta N/2)}; \beta = 1 (COE), = 2 (CUE), = 4 (CSE) \quad (4.2a)$$

The above two formulas, cited from Mehta's book [3], are now recapitulated from the Gaussian-reservoir point of view. Namely, the factor of multiple pair-product in the N -level joint probability density function (4.1) and (4.2) are just the reduced function of the form (3.12) so that

$$c(x - y) = (x - y)^2 \quad \text{for GUE} \quad (4.3)$$

and

$$c(x - y) = |e^{i(x-y)} - 1|^2 = [2\sin(x - y)/2]^2 \quad \text{for CUE.} \quad (4.4)$$

We show that these are the consequence of Balian's program, specifically, *postulate A* about the Riemannian metric in the information quantity applied to matrix spaces. The reason for this can be found in the diagonalization process of a sample matrix H (hermitian) in the bilinear form K_H , or U (unitary) in K_U (a sesqui-linear form). For the hermitian case:

$$H \mapsto \mathcal{U}^* H \mathcal{U} = \text{diag}(x_1, \dots, x_n) \equiv H_D \quad \text{or,} \quad H = \mathcal{U} \text{diag}(x_1, \dots, x_n) \mathcal{U}^*. \quad (4.5)$$

Then,

$$\mathcal{U} dH_D \mathcal{U}^* + d\mathcal{U} H_D \mathcal{U}^* + \mathcal{U} H_D d\mathcal{U}^* = \mathcal{U} (dH_D + [\mathcal{U}^* d\mathcal{U}, H_D]) \mathcal{U}^* \quad (d\mathcal{U}^* \mathcal{U} = -\mathcal{U}^* d\mathcal{U}) \quad (4.6)$$

where

$$H_D \equiv \text{diag}(x_1, x_2, \dots, x_N), \quad \text{and} \quad dH_D = \text{diag}(dx_1, dx_2, \dots, dx_N) \quad (4.7)$$

in the same representation, and hence

$$[H_D, dH_D] = 0. \quad (4.8)$$

Inserting this into eq.(1.1), we obtain

$$ds^2 = \text{tr} dH_D^2 + \text{tr}[\mathcal{U}^* d\mathcal{U}, H_D]^2 + 2\text{tr}\mathcal{U}^* d\mathcal{U}[H_D, dH_D] = \text{tr} dH_D^2 + \text{tr}[\mathcal{U}^* d\mathcal{U}, H_D]^2.$$

Hence,

$$ds^2 = \text{tr} dH_D^2 + \text{tr}[H_D, \mathcal{U}^* d\mathcal{U}]^2. \quad (4.9)$$

Accordingly, we have two ways of representing the metric of dH ; namely, in the original fixed representation (called *fixed frame*) and the H -diagonal representation (called *moving frame*). For the hermitian case,

$$ds^2 = \sum_n (dH_{nn})^2 + 2 \sum_{m < n} |dH_{mn}|^2 \quad (\text{fixed frame}) \quad (4.10)$$

$$= \sum_{j=1} dx_j^2 + 2 \sum_{j < k} (x_j - x_k)^2 |\Omega_{jk}|^2 \quad (\text{moving frame}) \quad (4.11)$$

where

$$\Omega \equiv \mathcal{U}^* d\mathcal{U} = -d\mathcal{U}^* \mathcal{U}. \quad (4.12)$$

Similarly, for the unitary case (which may reduce to the hermitian case by setting $U = U_0 e^{iH}$),

$$ds^2 = \sum_n (dU_{nn})^2 + 2 \sum_{m < n} |dU_{mn}|^2 \quad (\text{fixed frame}) \quad (4.13)$$

$$= \sum_{i=1} dx_i^2 + 2 \sum_{i < j} |e^{i(x_j - x_k)} - 1|^2 |\Omega_{ij}|^2 \quad (\text{moving frame}) \quad (4.14)$$

where Ω is the same as in (4.12). Therefore, by identifying $2|\Omega_{jk}|^2$ with the square of the tangent vector component $|A_{jk}|^2$, we get the expressions of the c-function eqs.(4.3,4).

5. Hamiltonian Formalism and Extended Distribution

A Riemannian metric form, quadratic with respect to a tangent vector, defines a classical mechanics, and the foregoing presentations (4.11) and (4.14) are adapted to such an expectation which may be called "level dynamics". Its motivation to the present Gaussian-reservoir distribution should be that the random variable that corresponds to the *cotangent vector* is expected to have a conjugate nature to velocity variables: this may be illustrated by the following chart:

$$\begin{aligned} \text{tangent vector} &\longrightarrow \text{cotangent vector} \\ \text{velocity} &\longrightarrow \text{momentum} \\ \text{angular velocity} &\longrightarrow \text{angular momentum} \\ \text{Lagrangian description} &\longrightarrow \text{Hamiltonian description} \end{aligned}$$

We outline the framework more precisely.

5.1. Conjugate variables; an adequate reservoir system

$$T = \frac{1}{2} \sum_j \left(\frac{dx_j}{dt} \right)^2 + \sum_{j < k} (x_j - x_k)^2 |A_{jk}|^2 \quad A_{jk} \equiv \Omega_{jk}/dt \quad (\text{angular velocity}) \quad (5.1)$$

for a hermitian level dynamics, and

$$T = \frac{1}{2} \sum_j \left(\frac{dx_j}{dt} \right)^2 + \sum_{j < k} |e^{i(x_j - x_k)} - 1|^2 |A_{jk}|^2 \quad A_{jk} \equiv \Omega_{jk}/dt \quad (\text{same as above}) \quad (5.2)$$

for a unitary level dynamics. (**remark.** The relation between $|\Omega|^2$ and $|A|^2$ is factor 2 different from the preceding section for the purpose of introducing Poisson brackets.)
conjugate momentum and angular momentum For the hermitian case

$$p_j = \frac{dx_j}{dt}, \quad M_{jk} = -2(x_j - x_k)^2 A_{jk}, \quad (5.3)$$

and for the unitary case

$$p_j = \frac{dx_j}{dt}, \quad M_{jk} = -2[2\sin(x_j - x_k)/2]^2 A_{jk}. \quad (5.4)$$

We have made a natural introduction of the angular momentum M as the conjugate momentum to the angular velocity: in matrix form

$$M = -2[H_D, [H_D, A]] \quad (\text{for the hermitian case}) \quad (5.5)$$

$$= -2[U_D^*, [U_D, A]] \quad (\text{for the unitary case}). \quad (5.6)$$

It can be shown that the components of M satisfies Poisson brackets relations as regards the Lie algebra for N -dimensional (real or complex) rotations[7] (see below). One can adopt another definition of the angular momentum by omitting the factor 2, thus

$$F = \frac{1}{2}M = -[H_D, [H_D, A]] \quad \text{or,} \quad -[U_D^*, [U_D, A]] \equiv (f_{jk});$$

$$f_{jk} = -(x_j - x_k)^2 A_{jk} \quad \text{or,} \quad -[2\sin(x_j - x_k)/2]^2 A_{jk}. \quad (5.7)$$

Hamiltonian function In terms of this form of the angular momentum $F = (f_{jk})$, we can write the level dynamics Hamiltonian function as

$$\mathcal{H} = \frac{1}{2} \sum_j p_j^2 + \frac{1}{2} \sum_{j \neq k} \frac{|f_{jk}|^2}{(x_j - x_k)^2} \quad (\text{the hermitian case}) \quad (5.8)$$

and

$$= \frac{1}{2} \sum_j p_j^2 + \frac{1}{2} \sum_{j \neq k} \frac{|f_{jk}|^2}{4\sin^2(x_j - x_k)/2} \quad (\text{the unitary case}). \quad (5.9)$$

We can now observe that the multi-dimensional angular momentum (f_{jk}) provides the relevant reservoir variable Y for the Gaussian distribution (3.8) with c -function (4.3) or (4.4). What is significant here is the Hamiltonian function so derived actually describes the level dynamics by means of equations of motion based on Poisson brackets(P.b.), as discussed next.

5.2. Origin of Poisson brackets

In general, $2f$ (f = the degree of freedom) coordinates form a smooth manifold with exterior derivatives, called *symplectic manifold* \mathcal{M}^{2f} , where symplectic differential forms with symbol \wedge play a basic role.

symplectic 1-form $\omega^{(1)} = \sum p_\alpha dq_\alpha$

$$\omega^{(1)} = \text{Tr} V dH = \sum P_{nn} dH_{nn} + \sum_{m < n} P_{mn} dH_{mn} \quad (\text{fixed frame}) \quad (5.10)$$

$$= \sum p_i dx_i + \sum_{i < j} M_{ij} \Omega_{ij} \quad (\text{moving frame}) \quad (5.11)$$

symplectic 2-form $\omega^{(2)} = \sum dp_\alpha \wedge dq_\alpha = d\omega^{(1)}$

$$(dx \wedge dy = -dy \wedge dx, dx \wedge dx = 0, (dx \wedge dy) \wedge dz = dx \wedge (dy \wedge dz), d(dx) = 0)$$

This is the canonical 2-form by which the canonical Poisson bracket (P.b.) relations are given to the $2f$ coordinates, making them canonical coordinates. An important subject here is to establish the P.b. relations involving the f non-canonical coordinates $\{M_\alpha\}$, each M_α being conjugate to Ω_α . Our answer can be found in the following statement.

Proposition *Let \mathbf{G} be the adjoint representation of a Lie algebra on a matrix space, generated by the basis set $\{E_\alpha\}$ ($\Omega = \sum_\alpha \Omega_\alpha E_\alpha$) satisfying the structure relations*

$$[E_\alpha, E_\beta] = \sum_{\gamma=1}^f c_{\alpha,\beta}^\gamma E_\gamma, \quad \alpha, \beta = 1, \dots, f. \quad (5.12)$$

There exists an operation $\{, \}$ on the symplectic manifold \mathcal{M}^f ; $C^\infty(\mathcal{M}^f) \times C^\infty(\mathcal{M}^f) \longrightarrow C^\infty(\mathcal{M}^f)$, called Berezin bracket for F and $G \in C^\infty(\mathcal{M}^f)$, given by

$$\{F, G\} = - \sum_{\alpha,\beta,\gamma} c_{\alpha,\beta}^\gamma M_\gamma \frac{\partial F}{\partial M_\alpha} \frac{\partial G}{\partial M_\beta}, \quad \text{in particular, } \{M_\alpha, M_\beta\} = - \sum_\gamma c_{\alpha,\beta}^\gamma M_\gamma. \quad (5.13)$$

This generates another representation (called the *coadjoint representation*) of the algebra. The Berezin bracket[8] is an example of more general form of binary operations $\{, \}$ on \mathcal{M}^{2f} , $\omega^{\mu\nu} \partial_\mu F \partial_\nu G$ (in the tensor notation) with a skew symmetric, nondegenerate $2f \times 2f$ tensor $\omega^{\mu\nu}$;

$$(\omega^{\mu\nu}) = (\omega_{\mu\nu})^{-1} \quad \text{where } \omega_{\mu\nu} = \text{coefficient tensor of } \omega^{(2)}; \quad d\omega^{(2)} = 0 \text{ } (\omega^{(2)} \text{ closed}). \quad (5.14)$$

The closedness of $\omega^{(2)}$ is equivalent to the Jacobi identity satisfied by the operation. We outline **proof** of the above **Proposition** by computing $\omega^{(2)} = d\omega^{(1)}$ given by eq.(5.11) in the moving frame. A direct proof that the Berezin bracket satisfies the Jacobi identity is given in a monograph by Fomenko and Trofimov[8].

symplectic 2-form in the moving frame

$$d \left(\sum_j p_j dx_j + \sum_{\alpha=(j < k)} M_\alpha \Omega_\alpha \right) = \sum_j dp_j \wedge dx_j + \sum_\alpha (dM_\alpha \wedge \Omega_\alpha + M_\alpha d\Omega_\alpha). \quad (5.15)$$

But we have $d\Omega = -\Omega \wedge \Omega$ which is called Maurer-Cartan equation[7] as can be derived from the definition (4.12) i.e. $d\Omega = -\mathcal{U}^*(d\mathcal{U} \wedge \Omega) = -(\mathcal{U}^* d\mathcal{U}) \wedge \Omega = -\Omega \wedge \Omega$. Hence in matrix forms,

$$(\omega_{\mu\nu}) = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \oplus \begin{pmatrix} C & -I \\ I & 0 \end{pmatrix}$$

a direct sum of two symplectic matrices; first, $f(= N)$ dimensional elementary one for the vector (dx, dp) and second, $f(= \frac{N}{2}N(N-1))$ dimensional non-elementary one for the vector (Ω, dM) . The inverse of the above matrix yields

$$(\omega^{\mu\nu}) = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & I \\ -I & C \end{pmatrix}$$

providing the coefficients of the Berezin bracket from the nonvanishing C part: namely, by inserting $\Omega = \sum \Omega_\alpha E_\alpha$ into the matrix 2-form $\Omega \wedge \Omega$ to get $\text{tr} M \Omega \wedge \Omega = \frac{1}{2} \sum_{\alpha, \beta} \Omega_\alpha \wedge \Omega_\beta \text{tr} M [E_\alpha, E_\beta]$, and by virtue of eq.(5.12) the coefficient of $\Omega_\alpha \wedge \Omega_\beta$ in $\text{tr} M d\Omega$ is equal to $-\sum_\gamma c_{\alpha, \beta}^\gamma M_\gamma$, where $M = \sum M_\gamma E_\gamma$. Note that the Berezin bracket for (f) 's is just one half of that for (M) 's in coefficients: $c_{\alpha\beta}^\gamma \rightarrow \frac{1}{2}c_{\alpha\beta}^\gamma$, in eq.(5.13).

Expression (5.15) then yields all the necessary P.b.'s such that (a) (x, p) 's are canonical P.b., (b) (f) 's are Berezin brackets and, (c) P.b.'s between (x, p) 's and (f) 's vanish. Thus, the foregoing discussion now establishes what is the precise meaning of our object and reservoir variables:

(object system) $x = (x_1, x_2, \dots, x_N)$; eigenvalues of a hermitian H , or, those in $U = U_0 e^{iH}$.

(reservoir system 1) $p = (\dots \frac{dx_j}{dt} \dots)$. These two systems constitute a canonical system where the usual form of canonical P.b. relations hold.

(reservoir system 2) $f_{jk}(j \neq k)$; all the off-diagonal elements in H -diagonal representation, satisfying the angular-momentum P.b.'s (the Berezin bracket relations) within the system, but its mechanical independence from the object (and the reservoir 1) system holds:

$$\{(x, p), f_{jk}\} = 0. \quad (5.16)$$

The Hamiltonian level dynamics is the one specified by the equations of motion $\dot{\eta} = \{\eta, \mathcal{H}\}$, where η stands for x, p , and f [9]. Roughly speaking, for any matrix, expressed in the representation in which a given hermitian H is diagonalized, all those matrix elements which are off-diagonal with respect to this representation (i.e. those matrices which are non-commuting with H) act, in the sense of the above dynamics, as reservoir variables against the eigenvalue system of H .

5.3. Most general form of the reduced distribution for the object system

We now come to the question: what is the most general form of the (reduced) eigenvalue system distribution? We have described the answer elsewhere in detail [9],[10]. Namely, in terms of a real, non-negative parameter a ,

$$c(x-y) = \frac{(x-y)^2}{1 + a^{-2}(x-y)^2} \quad (\text{the hermitian case}) \quad (5.17)$$

$$c(x - y) = \frac{4\sin^2(x - y)/2}{1 + a^{-2}4\sin^2(x - y)/2} \quad (\text{the unitary case}). \quad (5.18)$$

These are the most general covariance function for a single reservoir variable $y_\alpha = \sqrt{2}f_\alpha$ under conditions (a) \sim (d') and (d'') in the Hamiltonian description (5.8) and (5.9): it answers the question about the most general type of Riemannian metrics allowed by the level dynamics, as on the following reasoning.

- a. An equilibrium distribution of the Hamiltonian system is canonical type with candidates of every constant of motion in the exponential, and the search for the maximum entropy may be reduced to that for possible quadratic constants of the angular momentum (f_α).
- b. There exist only two quadratic constants of the above nature i.e. the Hamiltonian function and the Identity form as regards (f_α).

6. A Concrete Example

The best illustration of the argument in the preceding section would be provided by way of examples, and for this purpose let us see again an old paper; *A Family of One-Parameter Ensembles of Unitaries* by Gaudin[11]. It is worthwhile to cite its Abstract:

A one-parameter family of unitary matrix ensembles is studied. We define the ensemble $E_2(z)$ of unitary random matrices, whose eigenvalues are $\epsilon_j = \exp(2\pi i\varphi_j)$, by the following joint probability density:

$$P(\epsilon_1\epsilon_2\ldots\epsilon_n)d\varphi_1d\varphi_2\ldots d\varphi_n \propto \prod_{j<k} \left| \frac{\epsilon_j - \epsilon_k}{\epsilon_j - z\epsilon_k} \right|^2 d\varphi_1d\varphi_2\ldots d\varphi_n, \quad 0 \leq \varphi \leq 1. \quad (6.1)$$

It realizes a continuous interpolation between the distribution of the eigenvalues in the Dyson unitary ensemble E_2 for $z = 0$ and the uniform distribution of n random points on the unit circle for $z = 1$. The thermodynamic analogy with a circular or linear classical repulsive gas at temperature $\beta^{-1} = 1/2$ is developed. The isotherm $\beta = 2$ and the corresponding virial series are exactly calculated. All the correlation functions are given in the limit of an infinite linear gas or of an infinite series of levels. This model shows the short-range repulsion effect between eigenvalues but no long-range crystalline order, which is a strong characteristic of all ensembles so far studied.

For our present purpose, we show that Gaudin's model (6.1) is exactly the parameter a dependent distribution with covariance function (5.18): this is shown as follows. Noting that

$$|e^{i2\pi\varphi} - 1|^2 = 4\sin^2 2\pi\varphi/2 \quad \text{and} \quad |e^{i2\pi\varphi} - z|^2 = 4z(1 + \frac{1}{4z}(1 - z)^2\sin^2\pi\varphi),$$

we can write the right hand side of (6.1) as $d\varphi_1d\varphi_2\ldots d\varphi_n \times \prod_{j<k} c(\varphi_j - \varphi_k)$, where

$$c(\varphi_j - \varphi_k) = \frac{\sin^2\pi(\varphi_j - \varphi_k)}{\sinh^2\alpha + \sin^2\pi(\varphi_j - \varphi_k)}, \quad (6.2)$$

and thus, in eq.(5.18),

$$a^2 = \sinh^2\alpha, \quad \text{with} \quad z = e^{-2\alpha} \quad \text{in} \quad (6.1). \quad (6.2a)$$

Therefore, the model offers an exact interpolation scheme between the uniform distribution ($a = 0$) corresponding to Poisson statistics and that of the circular unitary ensemble ($a \rightarrow \infty$, if properly normalized, CUE). This scheme of Poisson-circular interpolation by means of parameter a is expected to hold for the other classes (i.e. COE and CSE), if the index 2 is replaced by $\beta(\nu$, in the present notation) = 1 and 4 in the general expression (3.12).

Furthermore, as Gaudin stated in the above **Abstract**, by means of a scaling of the angular variable φ to x such that

$$\varphi = \frac{x}{L}, \quad 0 \leq x \leq L, \quad (6.3)$$

and taking an appropriate limit

$$L \rightarrow \infty \quad \text{with} \quad n/L = \text{fixed}, \quad (6.4)$$

called *thermodynamic limit*, the model becomes a linear gas model with short-range repulsion between the gas atoms, where the variance function $c(x - y)$ reduces to (5.17).

Some significant results contained in Gaudin's paper are listed.

1. Exact form of the partition function

$$Z_n(\alpha) = \int_0^1 \dots \int_0^1 d\varphi_1 \dots d\varphi_n \prod_{j < k} \frac{\sin^2 \pi(\varphi_j - \varphi_k)}{\sinh^2 \alpha + \sin^2 \pi(\varphi_j - \varphi_k)} = n! e^{-n(n-1)\alpha} \prod_{k=1}^n \frac{1 - e^{-2\alpha}}{1 - e^{-2k\alpha}}.$$

2. Explicit expression of the 2-point correlation (and cluster) function

$$r_2(s) = 1 - \left| \int_0^1 dx \exp\left[i \frac{s}{a} \log(e^{2\pi a x} - 1)\right] \right|^2 \quad (Y_2(s) = \left| \int_0^1 dx \exp\left[i \frac{s}{a} \log(e^{2\pi a x} - 1)\right] \right|^2).$$

($\rho \equiv \lim n/L = 1$, i.e. the unfolding scale is used.)

3. Explicit expression of the 2-point form factor (Fourier transform of $Y_2(s)$)

$$b(k) = \frac{1}{2 \sinh(|k|a/2)} \left(\frac{e^{|k|a/2}}{2\pi a} \log[1 + e^{-|k|a}(e^{2\pi a} - 1)] - e^{-|k|a/2} \right).$$

These expressions enable us to get several useful results [10]. Among them, we exhibit here the information quantities $H[X, Y]$, $H[Y|X]$ and the entropy $H(X)$ (per atom) of the linear gas of Gaudin as follows.

free energy From the partition function 1,

$$-\frac{\beta}{n} F = \frac{1}{n} \log n! - (n-1)\alpha + \log(1 - e^{-2\alpha}) - \frac{1}{n} \sum_{k=1}^n \log(1 - e^{-2k\alpha}), \quad \beta = 2.$$

For $n \gg 1$, the summation is replaced by an integration.

energy We use the formula for the average of a quantity $\frac{1}{n} \sum_{j,k} f(x_j - x_k)$ by means of the two-point form factor $b(k)$ and the Fourier transform of f ; $\mathcal{F}_f(k)$, as follows.

$$\lim_{n \rightarrow \infty} \langle \frac{1}{n} \sum_{j \neq k} f(x_j - x_k) \rangle = \mathcal{F}_f(0) - \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}_f(k) b(k) dk$$

We show the average of the single-pair (potential) energy function, $V(x-y)$ which is given by $\frac{1}{2} \log(1 + \frac{a^2}{(x-y)^2})$, and

$$\mathcal{F}_V(k) = \frac{1}{2} \int_{-\infty}^{\infty} \log(1 + a^2/x^2) e^{ikx} dx = \frac{\pi(1 - e^{-a|k|})}{|k|}.$$

This together with the two-point form factor 3 now yield the necessary ingredients.

results

$$-\frac{H(X,Y)}{n} = \frac{2}{n} F = \pi a + 1 - \log \frac{2\pi a}{1 - e^{-2\pi a}} - \frac{1}{\pi a} \int_0^{\pi a} \frac{2k dk}{e^{2k} - 1}. \quad (6.5)$$

$$-\frac{H(Y|X)}{n} = \frac{2}{n} \langle U \rangle = \pi a - \int_0^{\infty} \frac{dk}{k} \left\{ \frac{2}{2\pi a} \log[1 + (e^{2\pi a} - 1)e^{-k}] - e^{-k} \right\}. \quad (6.6)$$

$$\begin{aligned} \frac{H(X)}{n} &= \log \left(\frac{2\pi a}{1 - e^{-2\pi a}} \right) - 1 \\ &+ \frac{1}{\pi a} \int_0^{\pi a} \frac{2k dk}{e^{2k} - 1} - \int_0^{\infty} \left(\frac{1}{2\pi a} \log[1 + (e^{2\pi a} - 1)e^{-k}] - e^{-k} \right) \frac{dk}{k}. \end{aligned} \quad (6.7)$$

7. Concluding Remark

- A. Fig.1 shows the quantity $\frac{1}{n} H(X)$ vs a , representing a degree of irregularity contained in a random sequence of eigenvalues(per level):its negative value indicates an information gain by the sequence compared with the fully uncorrelated (Poisson) sequence. Dyson[12] computed this quantity in the pure CU ($a \rightarrow \infty$) limit, showing the result $\frac{1}{n} H(X) = -\gamma$ (Euler's constant) -0.577 . Formula (6.7) agrees with this prediction.
- B. **statistical independence and identical distribution** of all the matrix units are the two strictly retained conditions in the present framework of random matrix distributions, and it is not strange that no long-range crystalline order of level systems is predicted by this framework, as Gaudin stated. In the recent development of random matrix theories for application to mesoscopic physics[2], possible modifications should be necessary in order to describe "localization phenomena", a big current interest.

References

1. R. Balian, Random Matrices and Information Theory, *Il Nuovo Cimento* **B57**,184(1968).
2. T. Guhr, A. Müller-Groeling, and H. A. Weidenmüller, Random Matrix Theories in Quantum Physics: Common Concepts, preprint for Physics Reports, to be published.
3. M. L. Mehta, *Random Matrices* (Academic, New York, 1991), 2nd ed.
4. Section 2 is based on the beginning part of text books for Information Theory, mainly a Japanese monograph, *Stochastic Processes and Entropy* by S. Ihara (Iwanami Syoten, 1984).

5. D. Petz, Monotone metrics on matrix spaces, *Linear Algebra and Applications* **244**,81 (1996).
6. S. Amari, *Differential-Geometrical Methodds in Statistics*, Lecture Notes in Statistics, **28**(1985); for Quantum Statistics, see D. Petz and Cs.Sudár, Geometry of quantum states, *J. Math. Phys.***37**(1996),2662.
7. D. H. Sattinger and O. L. Weaver, *Lie Groups and Algebras with Application to Physics, Geometry and Mechanics*, (1986).
8. A. T. Fomenko and V. V. Trofimov, *Integrable Systems on Lie Algebras and Symplectic Spaces*, (Gordon Breah Publishers, 1988).
9. H. Hasegawa, Dynamical Formulation of Quantum Level Statistics, *Open Sys. and Information Dyn.* **4**,350(1997).
10. H. Hasegawa and J. -Z. Ma, Intermediate level statistics with one-parameter random matrix ensembles, *J. Math. Phys.* 1998 *to be published*.
11. M. Gaudin, A one-parameter family of unitary ensembles, *Nuclear Physics*, **85**,545(1996).
12. F. J. Dyson, Statistical theory of energy levels of complex systems I, *J. Math. Phys.* **3**, 140(1962).

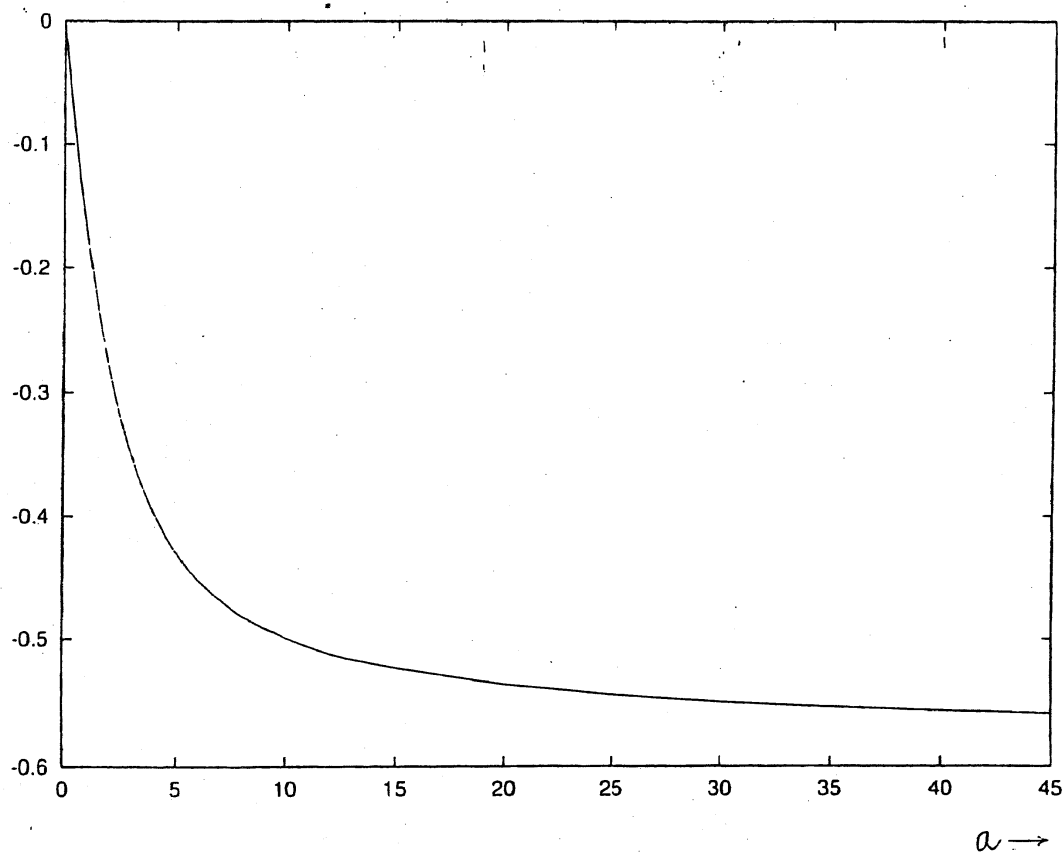


Fig.1 Information-loss curve for the Gaudin model of Poisson-CU ensembles.